"Express Mail" Mailing Label Number EL085318392US
Date of Deposit: May 19, 1999
I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above addressed to the Commissioner of Patents and Trademarks, Washington, D.C. 20231.
KURT FEUERSTEIN
(Typed or printed name of Sender)
(Signature)

OUTER MEMBRANE PROTEIN OF EHRLICHIA CANIS AND EHRLICHIA CHAFFEENSIS

This work was supported by grant RO1 AI33123 and RO1 AI40934 from National Institutes of Health. The government has certain rights in this invention.

BACKGROUND OF THE INVENTION

The ehrlichiae are obligate intracellular bacteria that infect circulating leucocytes. *Ehrlichia chaffeensis* infects the monocytes and macrophages in humans and causes human monocytic ehrlichiosis. The clinical manifestations of ehrlichiosis in humans are nonspecific and similar to Rocky Mountain spotted fever. The clinical manifestations include fever, chills, headache, myalgia or vomiting, and weight loss. Most patients have a history of tick exposure.

Ehrlichia canis infects and causes ehrlichiosis in animals belonging to the family Canidae. Canine ehrlichiosis consists of an acute and a chronic phase. The acute phase is characterized by fever, serous nasal and ocular discharges, anorexia, depression, and loss of weight. The chronic phase is characterized by severe pancytopenia, epistaxis, hematuria, blood in feces in addition to more severe clinical signs of the acute disease. If treated early during the course of the disease, dogs respond well to doxycycline. However, chronically infected dogs do not respond well to the antibiotic. Therefore, early diagnosis is very important for treating canine ehrlichiosis.

The primary diagnostic test for diagnosing canine ehrlichiosis and human ehrlichiosis is the indirect fluorescent antibody (IFA) test. This test uses the etiologic agent *Ehrlichia canis* to diagnose canine ehrlichiosis. The IFA test uses *Ehrlichia chaffeensis* as antigen for diagnosing human ehrlichiosis. The IFA test has, however, serious limitations. The IFA test is subject to false positives because the antigens are made of whole infected cells which comprise many nonspecific proteins which will cross-react with sera from some patients. The IFA test is also subject to false negatives because IFA antigens are unstable and may become inactivated during storage. In addition the IFA test requires a special equipment to perform the test. For example,

the IFA test requires a tissue culture system for growing the bacterium that are used to prepare the antigen slides, a fluorescent microscope, and trained persons to evaluate the serum reactivity to the bacterial antigen on the slide.

Tools which permit simpler, more rapid, and objective serodiagnosis of canine ehrlichiosis or human ehrlichiosis are desirable.

SUMMARY OF THE INVENTION

The present invention relates to improved diagnostic tools for veterinary and human use which are used for serodiagnosing ehrlichiosis in mammals, particularly in members of the Canidae family and in humans. The diagnostic tools are a group of outer membrane proteins of *E. chaffeensis* and variants thereof, referred to hereinafter as the "OMP proteins", a group of outer membrane proteins of *E. canis* and variants thereof referred to hereinafter as the "P30F proteins" , and antibodies to the OMP proteins and the P30F proteins.

The OMP proteins of *E. chaffeensis* encompass OMP-1, OMP-1A, OMP1-B, OMP-1C, OMP1-D, OMP1-E, OMP1-F, OMP1- H, OMP-1R, OMP-1S, OMP-1T, OMP -1U, OMP-1V, OMP-1W, OMP-1X, OMP-1Y and OMP-1Z. The mature OMP-1 protein of *E. chaffeensis* has a molecular weight of about 27.7 kDa and comprises amino acid 26 through amino acid 281 of the sequence shown in FIG.3B, SEQ ID NO: 2. The mature OMP-1B protein of *E. chaffeensis* has a molecular weight of about 28.2 kDa and comprises amino acid 26 through amino acid 283 of the sequence shown in FIG. 4B, SEQ ID NO: 4. The mature OMP-1C protein of *E. chaffeensis* has a molecular weight of about 27.6 kDa and comprises amino acid 26 through amino acid 280 of the sequence shown in FIG. 5B, SEQ ID NO: 6. The mature OMP-1D protein of *E. chaffeensis* has a molecular weight of about 28.7 and comprises amino acid 26 through amino acid 286 of the sequence shown in FIG. 6B, SEQ ID NO: 8. The mature OMP-1E protein of *E. chaffeensis* has a molecular weight of about 27.8 kDa and comprises amino acid 26 through amino acid 278 of the sequence shown in FIG. 7B, SEQ ID NO: 10. The mature OMP-1F protein of *E. chaffeensis* has a molecular weight of about 27.9 kDa and comprises amino acid 26 through amino acid 280 of the sequence shown in FIG. 8B, SEQ ID NO: 12. The mature OMP-1A protein of *E. chaffeensis* has a molecular weight of about 29.6 kDa and comprises amino acid 31 through amino acid 297 of the sequence shown in FIG. 9B, SEQ ID NO: 14. The mature OMP-1R protein of *E. chaffeensis* has a molecular weight of about 19.7 kDa and comprises amino acid 29 through amino acid 196 of the sequence shown in FIG. 10B, SEQ ID NO: 16. The mature OMP-1S protein of *E. chaffeensis* has a molecular weight of about 29.2 kDa and comprises amino acid 26 through amino acid 291 of the sequence shown in FIG. 11B, SEQ ID NO: 18. The OMP-1T protein of *E. chaffeensis* comprises the amino acid sequence shown in FIG. 12B, SEQ ID NO: 20. The mature

OMP-1U protein of *E. chaffeensis* has a molecular weight of about 30.6 kDa and comprises amino acid 26 through amino acid 295 of the sequence shown in FIG. 13B, SEQ ID NO: 22. The mature OMP-1V protein of *E. chaffeensis* has a molecular weight of about 28.0 kD and comprises amino acid 27 through amino acid 279 shown in FIG. 14B, SEQ ID NO: 24. The mature OMP-1W protein of *E. chaffeensis* has a molecular weight of about 28.8 kDa and comprises amino acid 30 through amino acid 283 of the sequence shown in FIG. 15B, SEQ ID NO: 26. The mature OMP-1X protein of *E. chaffeensis* has a molecular weight of about 27.8 kDa and comprises amino acid 25 through amino acid 275 of the sequence shown in FIG. 16B, SEQ ID NO: 28. The mature OMP-1Y protein of *E. chaffeensis* has a molecular weight about 28.8 kDa and comprises amino acid 28 through amino acid 285 of the sequence shown in FIG. 17B, SEQ ID NO: 30. The mature OMP-1Z protein of *E. chaffeensis* has a molecular weight of about 30.2 kDa and comprises amino acid 27 through amino acid 300 of the sequence shown in FIG. 18B, SEQ ID NO: 50. The mature OMP-1H protein has a molecular weight of about 30.2 kDa and comprises the amino acid 27 through amino acid 298 of sequence shown in Fig 33B, SEQ ID NO: 52.

The outer membrane proteins from *E. chaffeensis*, particularly a recombinant form of OMP-1, are immunogenic and, thus are useful for preparing antibodies. Such antibodies are useful for immunolabeling isolates of *E. chaffeensis* and for detecting the presence of *E. chaffeensis* in body fluids, tissues, and particularly in monocytes and macrophages. The OMP proteins, particularly OMP-1, are also useful for detecting antibodies to *E. chaffeensis* in the blood of patients with clinical signs of ehrlichiosis. The OMP protein, particularly OMP-1, are also useful immunogens for raising antibodies that are capable of reducing the level of infection in an immunized mammal that has been infected with *E. chaffeensis*. The proteins are also useful in a vaccine for protecting against infection with *E. chaffeensis*.

The P30F proteins of *E. canis* encompass P30, P30a, P30-1, P30-2, P30-3, P30-4, P30-5, P30-6, P30-7, P30-8, P30-9, P30-10, P30-11, and P30-12. The mature P30 protein of *E. canis* has a molecular weight of about 28.8 kDa and comprises amino acid 26 through amino acid 288 of the sequence shown in FIG. 19 B, SEQ ID NO: 32. The mature P30a protein of *E. canis* has a molecular weight of about 29.0 kDa and comprises amino acid 26 through amino acid 287 of the sequence shown in FIG. 20 B, SEQ ID NO: 34. The mature P30-1 protein of *E. canis* has a molecular weight of about 27.7 kDa and comprises amino acid 55 through amino acid 307 of the sequence shown in FIG. 21B, SEQ ID NO: 36. The mature P30-2 protein of *E. canis* has a molecular weight of about 28.0 kDa and comprises amino acid 26 through amino acid 280 of the sequence shown in FIG. 22 B, SEQ ID NO: 38. The mature P30-3 protein of *E. canis* has a

molecular weight of about 28.7 kDa and comprises amino acid 26 through amino acid 283 of the sequence shown in FIG. 23B, SEQ ID NO: 40. The mature P30-4 protein of *E. canis* has a molecular weight of about 28.0 kDa and comprises amino acid 26 through amino acid 276 of the sequence shown in FIG.24 B, SEQ ID NO: 42. The mature P30-5 protein of *E. canis* has a molecular weight of about 29.4 kDa and comprises amino acid 27 through amino acid 293 of the sequence shown in FIG. 25B, SEQ ID NO: 44. The mature P30-6 protein of *E. canis* has a molecular weight of about 29.4 kDa and comprises amino acid 31 through amino acid 293 of the sequence shown in FIG. 26B, SEQ ID NO: 54. The mature P30-7 protein of *E. canis* has a molecular weight of about 29.9 kDa and comprises amino acid 31 through amino acid 296 of the sequence shown in FIG. 27B, SEQ ID NO: 56. The mature P30-8 protein of *E. canis* has a molecular weight of about 30.3 kDa and comprises amino acid 27 through amino acid 299 of the sequence shown in FIG. 28 B, SEQ ID NO: 46. The mature P30-9 protein of *E. canis* has a molecular weight of about 28.6 kDa and comprises amino acid 27 through amino acid 281 of the sequence shown in FIG. 29B, SEQ ID NO: 58. The mature P30-10 protein of *E. canis* has a molecular weight of about 28.1 kDa and comprises amino acid 26 through amino acid 280 of the sequence shown in FIG. 30B, SEQ ID NO: 48. The mature P30-11 protein of *E. canis* has a molecular weight of about 28.6 kDa and comprises the amino acid 26 through amino acid 279 of sequence shown in FIG. 31B, SEQ ID NO: 60. The P30-12 protein of *E. canis* has a molecular weight of at least 27.3 kDa and comprises the amino acid sequence shown in FIG. 32B, SEQ ID NO: 62.

The P30F proteins, particularly P30, are immunogenic and are, thus, useful for preparing antibodies that are useful for immunolabeling isolates of *E. canis*. The P30 protein is also useful for diagnosing canine ehrlichiosis in mammals, particularly in members of the family Canidae, most particularly in dogs and for diagnosing infections with *E. chaffeensis* in humans. The P30F proteins are also useful immunogens for raising antibodies that reduce the level of infection in an immunized mammal that has been infected with *E. canis*. The P30F protein are also useful in a vaccine for protecting animals against infection with *E. canis*.

The present invention also provides isolated polynucleotides that encode the *E. chaffeensis* OMP proteins and isolated polynucleotides that encode the *E. canis* P30F proteins. The present invention also relates to antibodies which are immunospecific for and bind to the OMP proteins and the P30F proteins. Such antibodies are useful for immunolabeling isolates of *E. chaffeensis* and *E. canis*. The present invention also relates to kits containing reagents for diagnosing human ehrlichiosis and canine ehrlichiosis and to immunogenic compositions containing one or more OMP proteins or P30F proteins.

Brief Description of the Figures

FIG. 1. shows the DNA sequence and the amino acid sequence encoded by the *E. chaffeensis (p28)* gene cloned in pCRII*p28*. The N-terminal amino acid sequence of native *OMP-1* protein (P28) determined chemically is underlined. Five amino acid residues at the N terminus of P28 which were not included in the *p28* gene, are indicated by boldface. Arrows indicate annealing positions of the primer pair designed for PCR.

FIG. 2. shows the restriction map of 6.3-kb genomic DNA including the *omp-1* gene copies in *E. chaffeensis*. The four DNA fragments were cloned from the genomic DNA (pPS2.6, pPS3.6, pEC2.6, and pEC3.6). A recombinant plasmid pPS2.6 has an overlapping sequence with that of pEC3.6. The closed boxes at the bottom show PCR-amplified fragments from the genomic DNA for confirmation of the overlapping area. Open boxes at the top indicate open reading frames (ORF) of *omp-1* gene copies with direction by arrows. Open boxes at the bottom show DNA fragments subcloned for DNA sequencing.

FIG. 3B shows one embodiment of the OMP-1 protein; FIG. 3A shows one embodiment of the OMP-1 polynucleotide.

FIG. 4B shows one embodiment of the OMP-1B protein, FIG. 4A shows one embodiment of the OMP-1B polynucleotide

FIG. 5A shows one embodiment of the OMP-1C polynucleotide; FIG 5B shows one embodiment of the OMP-1C protein.

FIG. 6B shows one embodiment of the OMP-1D protein; FIG. 6A shows one embodiment of the OMP-1D polynucleotide.

FIG. 7B shows one embodiment of the OMP-1E protein; FIG 7A shows one embodiment of the OMP-1E polynucleotide.

FIG. 8B shows one embodiment of the OMP-1F protein; FIG 8A shows one embodiment of the OMP-1F polynucleotide.

FIG. 9B shows one embodiment of the OMP-1A protein, FIG 9A shows one embodiment of the OMP-1A polynucleotide.

FIG. 10 B shows one embodiment of a portion of the OMP-1R protein, FIG 10A shows one embodiment of an OMP-1R polynucleotide encoding such polypeptide.

FIG. 11 B shows one embodiment of a portion of the OMP-1S protein, FIG 11A shows one embodiment of the OMP-1S polynucleotide encoding such polypeptide.

FIG. 12 B shows one embodiment of a portion of the OMP-1T protein, FIG 12A shows one embodiment of the OMP-1T polynucleotide encoding such polypeptide.

FIG. 13 B shows one embodiment of the OMP-1U protein, FIG 13A shows one embodiment of the OMP-1U polynucleotide.

FIG. 14 B shows one embodiment of the OMP-1V protein, FIG 14A shows one embodiment of the OMP-1V polynucleotide.

FIG. 15 B shows one embodiment of the OMP-1W protein, FIG 15A shows one embodiment of the OMP-1W polynucleotide.

FIG. 16 B shows one embodiment of the OMP-1X protein, FIG 16A shows one embodiment of the OMP-1X polynucleotide.

FIG. 17 B shows one embodiment of the OMP-1Y protein, FIG 17A shows one embodiment of the OMP-1Y polynucleotide.

FIG. 18 B shows one embodiment of the OMP-1Z protein, FIG 18A shows one embodiment of the OMP-1Z polynucleotide.

FIG. 19 B shows one embodiment of the P30 protein, FIG 19A shows one embodiment of the P30 polynucleotide.

FIG. 20 B shows one embodiment of the P30a protein, FIG 20A shows one embodiment of the p30a polynucleotide.

FIG. 21 B shows one embodiment of the P30-1 protein, FIG 21A shows one embodiment of the p30-1 polynucleotide.

FIG. 22 B shows one embodiment of the P30-2 protein, FIG 22 A shows one embodiment of the p30-2 polynucleotide.

FIG. 23 B shows one embodiment of the P30-3 protein, FIG 23 A shows one embodiment of the p30-3 polynucleotide.

FIG. 24 B shows one embodiment of the P30-4 protein, FIG 22 A shows one embodiment of the p30-4 polynucleotide.

FIG. 25 B shows one embodiment of the P30-5 protein, FIG 22 A shows one embodiment of the p30-5 polynucleotide.

FIG. 26 B shows one embodiment of the P30-6 protein, FIG 26 A shows one embodiment of the p30-6 polynucleotide.

FIG. 27 B shows one embodiment of the P30-7 protein, FIG 27 A shows one embodiment of the p30-7 polynucleotide.

FIG. 28 B shows one embodiment of the P30-8 protein, FIG 28 A shows one embodiment of the p30-8 polynucleotide.

FIG. 29 B shows one embodiment of a portion of the P30-9 protein, FIG 29 A shows one embodiment of the p30-9 polynucleotide.

FIG. 30 B shows one embodiment of a portion of the P30-10 protein, FIG 30 A shows one embodiment of the p30-10 polynucleotide encoding such protein.

FIG. 31 B shows one embodiment of a portion of the P30-11 protein, FIG 31 A shows one embodiment of the p30-11 polynucleotide.

FIG. 32 B shows one embodiment of a portion of the P30-12 protein, FIG 32 A shows one embodiment of the p30-12 polynucleotide.

FIG. 33 B shows one embodiment of a portion of the OMP-1H protein, FIG 33 A shows one embodiment of the OMP-1H polynucleotide.

FIG. 34 depicts the amino acid sequences alignment of six *E. chaffeensis* OMP-1s and *Cowdria ruminantium* MAP-1. Aligned positions of identical amino acids with OMP-1F are shown with dots. The sequence of *C. ruminantium* MAP-1 is from the report of Van Vliet et al (1994) Molecular cloning, sequence analysis, and expression of the gene encoding the immunodominant 32-kilodalton protein of *Cowdria ruminantium*. Infect. Immun. 62:1451-1456. Gaps indicated by dashes were introduced for optimal alignment of all proteins. Bars indicate semivariable region (SV) and three hypervariable regions (HV1, HV2, and HV3).

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a group of outer membrane proteins of *E. chaffeensis,* OMP proteins, and a group of outer membrane proteins of *E. canis*, the P30F proteins. The mature OMP-1 protein of *E. chaffeensis* has a molecular weight of about 27.7 kDa and comprises amino acid 26 through amino acid 281 of the sequence shown in FIG.3B, SEQ ID NO: 2. The mature OMP-1B protein of *E. chaffeensis* has a molecular weight of about 28.2 kDa and comprises amino acid 26 through amino acid 283 of the sequence shown in FIG. 4B, SEQ ID NO: 4. The mature OMP-1C protein of *E. chaffeensis* has a molecular weight of about 27.6 kDa and comprises amino acid 26 through amino acid 280 of the sequence shown in FIG. 5B, SEQ ID NO: 6. The mature OMP-1D protein of *E. chaffeensis* has a molecular weight of about 28.7 and comprises amino acid 26 through amino acid 286 of the sequence shown in FIG. 6B, SEQ ID NO: 8. The mature OMP-1E protein of *E. chaffeensis* has a molecular weight of about 27.8 kDa and comprises amino acid 26 through amino acid 278 of the sequence shown in FIG. 7B, SEQ ID NO: 10. The mature OMP-1F protein of *E. chaffeensis* has a molecular weight of about 27.9 kDa and comprises amino acid 26 through amino acid 280 of the sequence shown in FIG. 8B, SEQ ID NO: 12. The mature OMP-1A protein of *E. chaffeensis* has a molecular weight of about 29.6 kDa and comprises amino acid 31 through amino acid 279 of the sequence shown in FIG. 9B, SEQ ID NO: 14. The mature OMP-1R protein of *E. chaffeensis* has a molecular weight of about 19.7 kDa

and comprises the amino acid 29 through amino acid 196 of the sequence shown in FIG. 10B, SEQ ID NO: 16. The mature OMP-1S protein of *E. chaffeensis* has a molecular weight of about 29.2 kDa and comprises amino acid 26 through amino acid 291 of the sequence shown in FIG. 11B, SEQ ID NO: 18. The OMP-1T protein of *E. chaffeensis* comprises the amino acid sequence shown in FIG. 12B, SEQ ID NO: 20. The mature OMP-1U protein of *E. chaffeensis* has a molecular weight of about 30.6 kDa and comprises amino acid 26 through amino acid 295 of the sequence shown in FIG. 13B, SEQ ID NO: 22. The mature OMP-1V protein of *E. chaffeensis* has a molecular weight of about 28.0 kD and comprises amino acid 27 through amino acid 279 shown in FIG. 14B, SEQ ID NO: 24. The mature OMP-1W protein of *E. chaffeensis* has a molecular weight of about 28.8 kDa and comprises amino acid 30 through amino acid 283 of the sequence shown in FIG. 15B, SEQ ID NO: 26. The mature OMP-1X protein of *E. chaffeensis* has a molecular weight of about 27.8 kDa and comprises amino acid 25 through amino acid 275 of the sequence shown in FIG. 16B, SEQ ID NO: 28. The mature OMP-1Y protein of *E. chaffeensis* has a molecular weight about 28.8 kDa and comprises amino acid 28 through amino acid 285 of the sequence shown in FIG. 17B, SEQ ID NO: 30. The mature OMP-1Z protein of *E. chaffeensis* has a molecular weight of about 30.2 kDa and comprises amino acid 27 through amino acid 300 of the sequence shown in FIG. 18B, SEQ ID NO: 50. The mature OMP-1H protein has a molecular weight of about 30.2 kDa and comprises the amino acid 27 through amino acid 298 of sequence shown in Fig 33B, SEQ ID NO: 52.

The mature P30 protein of *E. canis* has a molecular weight of about 28.8 kDa and comprises amino acid 26 through amino acid 288 of the sequence shown in FIG. 19 B, SEQ ID NO: 32. The mature P30a protein of *E. canis* has a molecular weight of about 29.0 kDa and comprises amino acid 26 through amino acid 287 of the sequence shown in FIG. 20 B, SEQ ID NO: 34. The mature P30-1 protein of *E. canis* has a molecular weight of about 27.7 kDa and comprises amino acid 55 through amino acid 307 of the sequence shown in FIG. 21B, SEQ ID NO: 36. The mature P30-2 protein of *E. canis* has a molecular weight of about 28.0 kDa and comprises amino acid 26 through amino acid 280 of the sequence shown in FIG. 22 B, SEQ ID NO: 38. The mature P30-3 protein of *E. canis* has a molecular weight of about 28.7 kDa and comprises amino acid 26 through amino acid 283 of the sequence shown in FIG. 23B, SEQ ID NO: 40. The mature P30-4 protein of *E. canis* has a molecular weight of about 28.0 kDa and comprises amino acid 26 through amino acid 276 of the sequence shown in FIG.24 B, SEQ ID NO: 42. The mature P30-5 protein of *E. canis* has a molecular weight of about 29.4 kDa and comprises amino acid 27 through amino acid 293 of the sequence shown in FIG. 25B, SEQ ID NO: 44. The mature P30-6 protein of *E. canis* has a molecular weight of about 29.4 kDa and

comprises amino acid 31 through amino acid 293 of the sequence shown in FIG. 26B, SEQ ID NO: 54. The mature P30-7 protein of *E. canis* has a molecular weight of about 29.9 kDa and comprises amino acid 31 through amino acid 296 of the sequence shown in FIG. 27B, SEQ ID NO: 56. The mature P30-8 protein of *E. canis* has a molecular weight of about 30.3 kDa and comprises amino acid 27 through amino acid 299 of the sequence shown in FIG. 28 B, SEQ ID NO: 46. The mature P30-9 protein of *E. canis* has a molecular weight of about 28.6 kDa and comprises amino acid 27 through amino acid 281 of the sequence shown in FIG. 29B, SEQ ID NO: 58. The mature P30-10 protein of *E. canis* has a molecular weight of about 28.1 kDa and comprises amino acid 26 through amino acid 280 of the sequence shown in FIG. 30B, SEQ ID NO: 48. The mature P30-11 protein of *E. canis* has a molecular weight of about 28.6 kDa and comprises the amino acid 26 through amino acid 279 of sequence shown in FIG. 31B, SEQ ID NO: 60. The P30-12 protein of *E. canis* has a molecular weight of at least 27.3 kDa and comprises the amino acid sequence shown in FIG.32B, SEQ ID NO: 62.

The present invention also encompasses variants of the OMP proteins shown in Figs 3-18 and 33 and variants of the P30F proteins shown in Figs 19-32. A "variant" as used herein, refers to a protein whose amino acid sequence is similar to one the amino acid sequences shown in Figs 3-33, hereinafter referred to as the reference amino acid sequence, but does not have 100% identity with the respective reference sequence. The variant protein has an altered sequence in which one or more of the amino acids in the reference sequence is deleted or substituted, or one or more amino acids are inserted into the sequence of the reference amino acid sequence. As a result of the alterations, the variant protein has an amino acid sequence which is at least 95% identical to the reference sequence, preferably, at least 97% identical, more preferably at least 98% identical, most preferably at least 99% identical to the reference sequence. Variant sequences which are at least 95% identical have no more than 5 alterations, i.e. any combination of deletions, insertions or substitutions, per 100 amino acids of the reference sequence. Percent identity is determined by comparing the amino acid sequence of the variant with the reference sequence using MEGALIGN project in the DNA STAR program. Sequences are aligned for identity calculations using the method of the software basic local alignment search tool in the BLAST network service (the National Center for Biotechnology Information, Bethesda, MD) which employs the method of Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* 215, 403-410. Identities are calculated by the Align program (DNAstar, Inc.) In all cases, internal gaps and amino acid insertions in the candidate sequence as aligned are not ignored when making the identity calculation.

While it is possible to have nonconservative amino acid substitutions, it is preferred that the substitutions be conservative amino acid substitutions, in which the substituted amino acid has similar structural or chemical properties with the corresponding amino acid in the reference sequence. By way of example, conservative amino acid substitutions involve substitution of one aliphatic or hydrophobic amino acids, e.g. alanine, valine, leucine and isoleucine, with another; substitution of one hydroxyl-containing amino acid, e.g. serine and threonine, with another; substitution of one acidic residue, e.g. glutamic acid or aspartic acid, with another; replacement of one amide-containing residue, e.g. asparagine and glutamine, with another; replacement of one aromatic residue, e.g. phenylalanine and tyrosine, with another; replacement of one basic residue, e.g. lysine, arginine and histidine, with another; and replacement of one small amino acid, e.g., alanine, serine, threonine, methionine, and glycine, with another.

The alterations are designed not to abolish the immunoreactivity of the variant protein with antibodies that bind to the reference protein. Guidance in determining which amino acid residues may be substituted, inserted or deleted without abolishing such immunoreactivity of the variant protein are found using computer programs well known in the art, for example, DNASTAR software. A variant of the OMP-1 protein is set forth in SEQ ID NO: 67 where the alanine at position 280 is replaced with a valine.

The present invention also encompasses fusion proteins in which a tag or one or more amino acids, preferably from about 2 to 65 amino acids, more preferably from about 34 to about 62 amino acids are added to the amino or carboxy terminus of the amino acid sequence of an OMP protein, a P30F protein, or a variant of such protein. Typically, such additions are made to stabilize the resulting fusion protein or to simplify purification of an expressed recombinant form of the corresponding OMP protein, P30F protein or variant of such protein. Such tags are known in the art. Representative examples of such tags include sequences which encode a series of histidine residues, the Herpes simplex glycoprotein D, or glutathione S-transferase.

The present invention also encompasses OMP proteins and P30F proteins in which one or more amino acids, preferably no more than 10 amino acids, in the respective OMP protein or P30F are altered by posttranslation processes or synthetic methods. Examples of such modifications include, but are not limited to, acetylation, amidation, ADP-ribosylation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or a lipid, cross-linking gamma-carboxylation, glycosylation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, sulfation, and transfer-RNA mediated additions of amino acids to proteins such as arginylation and ubiquitination.

The OMP proteins, particularly a recombinant form of OMP-1, are immunogenic and, thus are useful for preparing antibodies. Such antibodies are useful for immunolabeling isolates of *E. chaffeensis* and for detecting the presence of *E. chaffeensis* in body fluids, tissues, and particularly in monocytes and macrophages. The OMP proteins, particularly OMP-1, are also useful for detecting antibodies to *E. chaffeensis* in the blood of patients with clinical signs of ehrlichiosis. The OMP proteins , particularly OMP-1, are also useful immunogens for raising antibodies that are capable of reducing the level of infection in an immunized mammal that has been infected with *E. chaffeensis*. The OMP proteins are also useful in a vaccine for protecting against infection with *E. chaffeensis*.

The P30F proteins, particularly recombinant forms of P30, are immunogenic and are, thus, useful for preparing antibodies that are useful for immunolabeling isolates of *E. canis*. The P30 protein is also useful for diagnosing canine ehrlichiosis in mammals, particularly in members of the family Canidae, most particularly in dogs and for diagnosing infections with *E. chaffeensis* in humans. The P30F proteins are also useful immunogens for raising antibodies that reduce the level of infection in an immunized mammal that has been infected with *E. canis*. The P30F proteins are also useful in a vaccine for protecting animals against infection with *E. canis*.

In another aspect, the present invention provides a polypeptide which comprises a fragment of the OMP1 protein, hereinafter referred to as "rOMP-1" . The rOMP-1 polypeptide weighs approximately 31 kDa and comprises all but of the first 5 amino acids of mature OMP-1 protein. The rOMP-1 polypeptide comprises the amino acid sequence extending from amino acid 6 through amino acid 251 of the amino acid sequence shown in Fig.1, SEQ ID NO. 2. The present invention also embraces polypeptides where one or more of the amino acids in the sequence extending from amino acid 1 or 6 through amino acid 251 Fig. 1 are replaced by conservative amino acid residues. The present invention also relates to variant of rOMP-1 that have an amino acid sequence identity of at least 95%, more preferably at least 97%, and most preferably of at least 99% with the amino acid sequence extending from amino acid 6 through amino acid 251 of the OMP-1 protein and which derivative binds to antibodies in sera from humans infected with *E. chaffeensis*.

Polynucleotides

The present invention also provides isolated polynucleotides which encode the OMP proteins and the P30F proteins. The OMP-1 polynucleotide encodes the OMP-1 protein of *E. chaffeensis*, Figure 3A shows one embodiment of the OMP-1 polynucleotide, SEQ ID NO: 1. The OMP-1B polynucleotide encodes the OMP-1B protein of *E. chaffeensis*; Figure 4A shows one embodiment of the OMP-1B polynucleotide, SEQ ID NO: 3. The OMP-1C polynucleotide

encodes the OMP-1C protein of *E. chaffeensis*, Figure 5A shows one embodiment of the OMP-1C polynucleotide; SEQ ID NO: 5. The OMP-1D polynucleotide encodes the OMP-1D protein of *E. chaffeensis*; Figure 6A shows one embodiment of the OMP-1D polynucleotide, SEQ ID NO: 7. The OMP-1E polynucleotide encodes the OMP-1E protein of *E. chaffeensis;* Figure 7A shows one embodiment of the OMP-1E polynucleotide, SEQ ID NO: 9. The OMP-1F polynucleotide encodes the OMP-1F protein of *E. chaffeensis*; Figure 8A shows one embodiment of the OMP-1F polynucleotide, SEQ ID NO: 11. The OMP-1A polynucleotide encodes the OMP-1A protein of *E. chaffeensis*; Figure 9A shows one embodiment of the OMP-1A polynucleotide, SEQ ID NO: 13. The OMP-1R polynucleotide encodes the OMP-1R protein, Figure 10A shows one embodiment of a portion of the OMP-1R polynucleotide, SEQ ID NO: 15. The OMP-1S polynucleotide encodes the OMP-1S protein of *E. chaffeensis*; Figure 11A shows one embodiment of a portion of the OMP-1S polynucleotide, SEQ ID NO: 17. The OMP-1T polynucleotide encodes the OMP-1T protein of *E. chaffeensis*; Figure 12A shows one embodiment of a portion of the OMP-1T polynucleotide, SEQ ID NO: 19. The OMP-1U polynucleotide encodes the OMP-1U protein of *E. chaffeensis*; Figure 13A shows one embodiment of the OMP-1U polynucleotide, SEQ ID NO: 21. The OMP-1V polynucleotide encodes the OMP-1V protein of *E. chaffeensis*; Figure 14A shows one embodiment of the OMP-1V polynucleotide, SEQ ID NO: 23. The OMP-1W polynucleotide encodes the OMP-1W protein of *E. chaffeensis*; Figure 15A shows one embodiment of the OMP-1W polynucleotide, SEQ ID NO: 25. The OMP-1X polynucleotide encodes an OMP-1X protein of *E. chaffeensis*; Figure 16A shows one embodiment of the OMP-1X polynucleotide, SEQ ID NO 27. The OMP-1Y polynucleotide encodes the OMP-1Y protein of *E. chaffeensis*; Figure 17A shows one embodiment of the OMP-1Y polynucleotide, SEQ ID NO 29. The OMP-1Z polynucleotide encodes the OMP-1Z protein of *E. chaffeensis*; Figure 18A shows one embodiment of an OMP-1Z polynucleotide encoding such polypeptide, SEQ ID NO: 49. The OMP-1H polynucleotide encodes the OMP-1H protein of *E. chaffeensis*; Figure 33A shows one embodiment of a portion of the OMP-1H polynucleotide, SEQ ID NO: 51.

The p30 polynucleotide encodes the P30 protein of *E. canis*, Figure 19A shows one embodiment of the p30 polynucleotide, SEQ ID NO: 31. The p30a polynucleotide encodes the P30a protein of *E. canis*, Figure 20A shows one embodiment of the p30a polynucleotide, SEQ ID NO: 33. The p30-1 polynucleotide encodes the P30-1 protein of *E. canis*; Figure 21A shows one embodiment of the p30-1 polynucleotide, SEQ ID NO: 35. The p30-2 polynucleotide encodes the P30-2 protein of *E.canis*; Figure 22A shows one embodiment of the p30-2 polynucleotide, SEQ ID NO: 37. The p30-3 polynucleotide encodes the P30-3 protein of *E.canis*; Figure 23A shows

one embodiment of the p30-3 polynucleotide, SEQ ID NO: 39. The p30-4 polynucleotide encodes the P30-4 protein of *E. canis*, Figure 24A shows one embodiment of the p30-4 polynucleotide, SEQ ID NO: 41. The p30-5 polynucleotide encodes the P30-5 protein of *E. canis*, Figure 25A shows one embodiment of the p30-5 polynucleotide, SEQ ID NO: 43. The p30-6 polynucleotide encodes the P30-6 protein, Figure 26A shows one embodiment of the p30-6 polynucleotide, SEQ ID NO: 53. The p30-7 polynucleotide encodes the P30-7 protein of *E. canis*; Figure 27A shows one embodiment of the p30-7 polynucleotide, SEQ ID NO: 55. The p30-8 polynucleotide encodes the P30-8 protein of *E. canis*; Figure 28A shows one embodiment of the p30-8 polynucleotide, SEQ ID NO: 45. The p30-9 polynucleotide encodes the P30-9 protein of *E. canis*; Figure 29A shows one embodiment of a portion of the p30-9 polynucleotide, SEQ ID NO: 57. The p30-10 polynucleotide encodes the P30-10 protein of *E. canis*, Figure 30A shows one embodiment of a portion of the p30-10 polynucleotide, SEQ ID NO: 47. The p30-11 polynucleotide encodes the P30-11 protein of *E. canis*; Figure 31A shows one embodiment of a portion of the p30-11 polynucleotide, SEQ ID NO: 59. The p30-12 polynucleotide encodes the P30-12 protein of *E. canis*; Figure 32A shows one embodiment of a portion of the p30-12 polynucleotide, SEQ ID NO: 61.

The polynucleotides are useful for producing the outer membrane proteins of *E. chaffeensis* and *E. canis*. For example, an RNA molecule encoding the outer membrane protein OMP-1 is used in a cell-free translation systems to prepare OMP-1. Alternatively, a DNA molecule encoding the outer membrane protein is introduced into an expression vector and used to transform cells. Suitable expression vectors include for example chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40, bacterial plasmids, phage DNAs; yeast plasmids, vectors derived from combinations of plasmids and phage DNAs, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. The DNA sequence is introduced into the expression vector by conventional procedures.

Accordingly, the present invention also relates to recombinant constructs comprising one or more of the polynucleotide sequences. Suitable constructs include, for example, vectors, such as a plasmid, phagemid, or viral vector, into which a sequence that encodes the outer membrane protein has been inserted. In the expression vector, the DNA sequence which encodes the outer membrane protein is operatively linked to an expression control sequence, i.e., a promoter, which directs mRNA synthesis. Representative examples of such promoters, include the LTR or SV40 promoter, the *E. coli* lac or trp, the phage lambda PL promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or in viruses. The promoter may also be the natural promoter of the outer membrane protein coding sequence. The expression

vector also contains a ribosome binding site for translation initiation and a transcription terminator. Preferably, the recombinant expression vectors also include an origin of replication and a selectable marker, such as for example, the ampicillin resistance gene of *E. coli* to permit selection of transformed cells, i.e. cells that are expressing the heterologous DNA sequences. The polynucleotide sequence encoding the outer membrane protein is incorporated into the vector in frame with translation initiation and termination sequences. Optionally, the sequence encodes a fusion outer membrane protein which includes an N-terminal or C-terminal peptide or tag that stabilizes or simplifies purification of the expressed recombinant product. Representative examples of such tags include sequences which encode a series of histidine residues, the Herpes simplex glycoprotein D, or glutathione S-transferase.

Polynucleotides encoding the OMP proteins and the P30F proteins are also useful for designing hybridization probes for isolating and identifying cDNA clones and genomic clones encoding the OMP proteins, the P30F proteins or allelic forms thereof. Such hybridization techniques are known to those of skill in the art. The sequences that encode the OMP proteins and the P30F proteins are also useful for designing primers for polymerase chain reaction (PCR), a technique useful for obtaining large quantities of cDNA molecules that encode the OMP proteins and the P30F proteins.

Also encompassed by the present invention, are single stranded polynucleotides, hereinafter referred to as antisense polynucleotides, having sequences which are complementary to the DNA and RNA sequences which encode the OMP proteins and the P30F proteins. The term complementary as used herein refers to the natural binding of the polynucleotides under permissive salt and temperature conditions by base pairing,

The present invention also encompasses oligonucleotides that are used as primers in polymerase chain reaction (PCR) technologies to amplify transcripts of the genes which encode the OMP proteins, the P30F proteins or portions of such transcripts. Preferably, the primers comprise 18-30 nucleotides, more preferably 19-25 nucleotides. Preferably, the primers have a G+C content of 40% or greater. Such oligonucleotides are at least 98% complementary with a portion of the DNA strand, i.e., the sense strand, which encodes the OMP protein or the P30F protein, or a portion of its corresponding antisense strand. Preferably, the primer has at least 99% complementarity, more preferably 100% complementarity, with such sense strand or its corresponding antisense strand. Primers which are which have 100% complementarity with the antisense strand of a double-stranded DNA molecule which encodes an OMP protein or a P30F protein have a sequence which is identical to a sequence contained within the sense strand. The identity of primers which are 15 nucleotides in length and have full complementarity with a

portion of the antisense strand of a double-stranded DNA molecule which encodes the OMP-1 protein is determined using the nucleotide sequence, SEQ ID NO:1 , shown in FIG 3A and described by the general formula a-b, where a is any integer between 1 to 843, where b is equal to a+14, and where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:1.

The present invention also encompasses oligonucleotides that are useful as hybridization probes for detecting transcripts of the genes which encode the OMP proteins and P30F proteins or for mapping of the genes which encode the OMP proteins and P30F proteins. Preferably, such oligonucleotides comprise at least 210 nucleotides, more preferably at least 230, most preferably from about 210 to 280 nucleotides. Such hybridization probes have a sequence which is at least 90% complementary with a sequence contained within the sense strand of a DNA molecule which encodes each of OMP proteins and P30F proteins or with a sequence contained within its corresponding antisense strand. Such hybridization probes bind to the sense strand under stringent conditions. The term "stringent conditions" as used herein is the binding which occurs within a range from about Tm 5°C (5°C below the melting temperature Tm of the probe) to about 20°C to 25°C below Tm. The probes are used in Northern assays to detect transcripts of OMP and P30F homologous genes and in Southern assays to detect OMP and P30F homologous genes. The identity of probes which are 200 nucleotides in length and have full complementarity with a portion of the antisense strand of a double-stranded DNA molecule which encodes the OMP-1 protein is determined using the nucleotide sequence, SEQ ID NO: 1, shown in FIG 3A and described by the general formula a-b, where a is any integer between 1 to 843, b is equal to a +200, and where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:1.

The present invention also encompasses isolated polynucleotides which are alleles of the genes which encode the OMP proteins and the P30F proteins. As used herein, an allele or allelic sequence is an alternative form of the gene which may result from one or more mutations in the sequences which encode the OMP proteins and P30F proteins. Such mutations typically arise from natural addition, deletion of substitution of nucleotides in the open reading frame sequences. Any gene may have none, one, or several allelic forms. Such alleles are identified using conventional techniques, such as for example screening libraries with probes having sequences identical to or complementary with one or more OMP or P30F polynucleotides.

The present invention also encompasses altered polynucleotides which encode OMP proteins and P30F proteins. Such alterations include deletions, additions, or substitutions. Such alterations may produce a silent change and result in an OMP protein or P30F protein having the

same amino acid sequence as the OMP protein or P30F protein encoded by the unaltered polynucleotide. Such alterations may produce a nucleotide sequence possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eucaryotic host may be incorporated into the nucleotide sequences shown in Figures 3-33 to increase the rate of expression of the proteins encoded by such sequences. Such alterations may also introduce new restriction sites into the sequence or result in the production of an OMP protein variant or P30F protein variant. Typically, such alterations are accomplished using site-directed mutagenesis.

Antibodies

In another aspect, the present invention relates to antibodies which are specific for and bind to at least one OMP protein or P30F protein. Such antibodies are useful research tools for identifying cells, particularly monocytes or macrophages, infected with *E. chaffeensis* or *E. canis* and for purifying the major outer membrane protein of *E. chaffeensis* or *E. canis* from partially purified preparations by affinity chromatography. Such antibodies are also useful for identifying bacterial colonies, particularly colonies of genetically-engineered bacteria, that are expressing the major outer membrane protein of *E. chaffeensis* or *E. canis*.

Kits

The present invention also relates to kits containing reagents for diagnosing *E. chaffeensis* and *E. canis*. The kit comprises one or more OMP proteins, or one or more *E. canis* proteins, or antigenic fragments thereof. For ease of detection, it is preferred that the OMP protein or P30F proteins be attached to a substrate such as a column, plastic dish, matrix, or membrane, preferably nitrocellulose. The kit may further comprise a biomolecule, preferably a secondary antibody, for detecting interactions between the isolated OMP protein or P30F protein and antibodies in a patient sample. Preferably, the biomolecule is coupled to a detectable tag such as an enzyme, chromophore, fluorophore, or radio-isotope. The kit is used by contacting a patient sample with the OMP protein or P30F protein under conditions that permit formation of antigen-antibody complexes. Then the biomolecule is added and the presence or absence of any resulting antigen-antibody complexes is detected by assaying for a change in the sample, for example, by observing the formation of a precipitate in the sample, the presence of radioactivity on the substrate, or a color change in the sample or on the substrate.

Diagnostic Method

The present invention also provides a method for detecting antibodies to the *E. chaffeensis* or *E. canis* in a sample of a bodily fluid from a patient. The method comprises providing an isolated outer membrane protein of *E. chaffeensis* or *E. canis*, particularly a recombinant form of the isolated protein, contacting the outer membrane protein or polypeptide

with a sample taken from the patient; and assaying for the formation of a complex between the outer membrane protein or polypeptide and antibodies in the sample. For ease of detection, it is preferred that the isolated protein or polypeptide be attached to a substrate such as a column, plastic dish, matrix, or membrane, preferably nitrocellulose. The sample may be a tissue or a biological fluid, including urine, whole blood, or exudate, preferably serum. The sample may be untreated, subjected to precipitation, fractionation, separation, or purification before combining with the isolated protein or peptide. Interactions between antibodies in the sample and the isolated protein or peptide are detected by radiometric, colorimetric, or fluorometric means, size-separation, or precipitation. Preferably, detection of the antibody-outer membrane protein complex is by addition of a secondary antibody that is coupled to a detectable tag, such as for example, an enzyme, fluorophore, or chromophore. Formation of the complex is indicative of the presence of anti-*E chaffeensis* or anti-*E canis* antibodies, either IgM or IgG, in the patient. Thus, the method is used to determine whether a patient is infected with *E. chaffeensis* or *E. canis*.

Preferably, the method employs an enzyme-linked immunosorbent assay (ELISA) or a Western immunoblot procedure . Such methods are relatively simple to perform and do not require special equipment as long as membrane strips are coated with a high quality antigen. Accordingly, it is more advantageous to use a recombinant form of the outer membrane protein of *E. chaffeensis* or *E. canis* since such proteins, typically, are more pure and consistent in quality than a purified form of such protein.

Immunogenic Composition

The present invention also relates to immunogenic compositions comprising one or more OMP protein of *E. chaffeensis* and a pharmaceutically acceptable adjuvant and to immunogenic compositions comprising one or more P30F proteins of *E. canis* and a pharmaceutically acceptable adjuvant, which, preferably, enhances the immunogenic activity of the outer membrane protein in the host animal.

Preparing the OMP proteins and the P30F proteins

The OMP proteins and P30F proteins may be produced by conventional peptide synthesizers. The OMP proteins and P30F proteins may also be produced using cell-free translation systems and RNA molecules derived from DNA constructs that encode the OMP proteins and P30F proteins. Alternatively, OMP proteins and P30F proteins are made by transfecting host cells with expression vectors that comprise a DNA sequence that encodes the respective OMP protein or P30F protein and then inducing expression of the protein in the host cells. For recombinant production, recombinant constructs comprising one or more of the

sequences which encode the OMP protein or P30F protein are introduced into host cells by conventional methods such as calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape lading, ballistic introduction or infection.

The OMP proteins or P30F proteins may be expressed in suitable host cells, such as for example, mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters using conventional techniques. Following transformation of the suitable host strain and growth of the host strain to an appropriate cell density, the cells are harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification of the OMP protein or P30F protein.

Conventional procedures for isolating recombinant proteins from transformed host cells, such as isolation by initial extraction from cell pellets or from cell culture medium, followed by salting-out, and one or more chromatography steps, including aqueous ion exchange chromatography, size exclusion chromatography steps, and high performance liquid chromatography (HPLC), and affinity chromatography may be used to isolate recombinant OMP protein or P30F protein

Preparation of Antibodies

The OMP proteins, P30F proteins, and variants thereof are used as immunogens to produce antibodies immunospecific for one or more OMP protein or one or more P30F protein. The term "immunospecific" means the antibodies have substantially greater affinity for one or more OMP protein or P30F protein than for other proteins. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, and Fab fragments.

Polyclonal antibodies are generated using conventional techniques by administering the OMP protein or P30F protein, or a chimeric molecule to a host animal. Depending on the host species, various adjuvants may be used to increase immunological response. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin, and Corynebacterium parvum are especially preferable. Conventional protocols are also used to collect blood from the immunized animals and to isolate the serum and or the IgG fraction from the blood.

For preparation of monoclonal antibodies, conventional hybridoma techniques are used. Such antibodies are produced by continuous cell lines in culture. Suitable techniques for preparing monoclonal antibodies include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV hybridoma technique.

Various immunoassays may be used for screening to identify antibodies having the desired specificity. These include protocols which involve competitive binding or

immunoradiometric assays and typically involve the measurement of complex formation between the respective OMP protein or P30F protein and the antibody.

Polynucleotides that encode OMP proteins and P30F proteins

Polynucleotides comprising sequences encoding an OMP protein or P30F protein may be synthesized in whole or in part using chemical methods. Polynucleotides which encode an OMP protein or P30F protein, particularly alleles of the genes which encode an OMP protein or P30F protein, may be obtained by screening a genomic library of an *E. chaffeensis* or *E. canis* isolate with a probe comprising sequences identical or complementary to the sequences shown in Figures 3-33 or with antibodies immunospecific for a OMP protein or P30F protein to identify clones containing such polynucleotide.

Polynucleotides which Encode OMP-1 protein and P30 protein

A. Isolation of the Outer Membrane Proteins

E. chaffeensis Arkansas strain and *E. canis* Oklahoma strain were cultivated in the DH82 dog macrophage cell line and purified by Percoll density gradient centrifugation. Purified ehrlichiae (100 μg) were suspended with 10 mM sodium phosphate buffer, pH 7.4, containing 0.1% Sodium *N*-lauroyl sarcosine (Sarkosyl) [Sigma, St. Louis, MO], 50 μg/ml each DNase I (Sigma) and RNase A (Sigma), and 2.5 mM $MgCl_2$. After incubation at 37° for 30 min, the sample was separated by centrifugation at 10,000 x g for 1 h into the soluble supernatant and the insoluble precipitate. The insoluble pellet was resuspended 2 to 3 times with 0.1% Sarkosyl and centrifuged. The final pellet was analyzed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and by electron microscopy.

Transmission electron microscopy revealed that the purified ehrlichial fraction consists of a mixture of electron dense and light forms of *E. chaffeensis* with slight disintegration of inner membrane . Ehrlichiae were not surrounded with the host inclusion membrane. Various sizes of membrane vesicles (< 1 μm) without significant ribosomes or nuclear materials were observed in the Sarkosyl-insoluble fraction from the organism. Succinic dehydrogenase (inner membrane marker enzyme of gram negative bacteria) activities were at less than the detection limit (1 n moles / min / mg of protein) in the Sarkosyl-insoluble fraction compared to approximately 10 n moles / min / mg of protein in the Percoll-purified organisms, suggesting that the insoluble fraction primarily consisted of the outer membrane of *E. chaffeensis.*

Analysis of the Sarkosyl-soluble, and insoluble fraction of *E. chaffeensis* by SDS-PAGE suggested that proteins of 30-kDa range in the insoluble fraction represent the major outer membrane proteins of this organism. Analysis of the Sarkosyl-soluble, and insoluble fraction of

E. canis by SDS-PAGE suggested that proteins of 30-kDa range in the insoluble fraction represent the major outer membrane proteins of this organism also. *E. canis* was antigenically cross reactive with *E. chaffeensis*. These findings indicate that the 30-kDa range proteins represent the major outer membrane proteins of these two *Ehrlichia* spp.

To improve resolution of the outer membrane proteins, proteins in the Sarkosyl-insoluble pellet prepared from 400 μg of purified *E. chaffeensis* were separated by a reversed-discontinuous (Rd) SDS-PAGE (2.5-cm-long 17% gel on top of 11-cm-long 12% gel). At least five proteins of 30-kDa range in *E. chaffeensis* (P23, P25, P27, P28, and P29) were resolved from the Sarkosyl-insoluble proteins.

B. Cloning and sequencing of the *omp-1* gene

The portion of the membrane containing bound proteins was excised and analyzed with an Applied Biosystems protein sequencer (Model 470). The N-terminal amino acid sequence of OMP-1 protein was determined as D P A G S G I N G N F Y I S G K Y M P, SEQ ID NO: 63. Based on 6th to 12th amino acids of this sequence, a forward primer, FECH1, having the sequence: 5'-CGGGATCCGAATTCGG(A/T/G/C)AT(A/T/C)AA(T/C)GG(A/T/G/C)AA(T/C)TT(T/C)TA-3'. SEQ ID NO:64 was designed. Amino acids at the 1 to 5 positions of the N terminus of OMP-1 were not included in this primer design. For insertion into an expression vector, a 14-bp sequence (underlined) was added at the 5' end of primer to create an *Eco*RI and a *Bam*HI site. The reverse primer, RECH2, which includes a *Not*I site at the 5' end for ligation into an expression vector had the sequence : 5'-AGCGGCCGCTTA(A/G)AA(T/C)A(C/G) (A/G)AA (C/T)CT T(C/G)C TCC-3'. SEQ ID NO:65.

Genomic DNA of *E. chaffeensis* was isolated from purified organisms. PCR amplification with FECH1 and RECH2 primers was performed using a Perkin-Elmer Cetus DNA Thermal Cycler (model 480). A 0.8-kb amplified product was cloned in the pCRII vector of a TA closing kit, as described by the manufacturer (Invitrogen Co., San Diego, CA). The clone obtained was designated pCRII*p28*. Both strands of the inserted DNA were sequenced by a dideoxy-termination method with an Applied Biosystems 373A DNA sequencer.

The 0.8-kb DNA fragment containing a partial OMP-1 gene, cloned in pCRII*p28*, had an open reading frame (ORF) of 756 bp encoding a 251-amino acid recombinant protein (including both PCR primer regions) with a molecular mass of 27.2 kDa. The nucleotide sequence of the open reading frame, and the amino acid sequence of the polypeptide of the partial OMP-1 protein, are shown in Fig. 1.

A DNA fragment comprising the partial *p30* gene was prepared in a similar manner, i.e., by PCR amplification of genomic DNA of *E. canis* using the forward primer, FECH1, which is described above, and a reverse primer, REC1, which is complimentary to the DNA sequence corresponding to amino acid positions 185 to 191 of the mature OMP-1 of *E. chaffeensis*. The sequence of REC1 is 5'-ACCTAACTTTCCTTGGTAAG-3', SEQ ID NO:66.

Genomic DNA of *E. canis* was isolated from the purified organism. PCR amplification was performed by using a Perkin-Elmer Cetus DNA Thermal Cycler (model 480). The 0.6-kb products were amplified with the FECH1-REC1 primer pair and were cloned into the pCRII vector of a TA cloning kit (Invitrogen Co., San Diego, CA). The clone obtained by the primer pair was designated pCRII*p30*. Both strands of the insert DNA were sequenced by a dideoxy termination method with an Applied Biosystems 373 DNA sequencer.

The 0.6-kb DNA fragment containing a partial *p30* gene cloned had an open reading frame (ORF) of 579 bp encoding a 193-amino-acid protein with a molecular mass of 21,175 Da. The partial P30 protein of *E. canis* was encoded by nucleotide 97 through nucleotide 672 of the sequence shown in Fig. 19A and comprised amino acid 33 through amino acid 224 of the sequence shown in Fig. 19B.

Polynucleotides which encode OMP 1A, OMP-1B, OMP-1C, OMP-1D, OMP-1F, and OMP1-E

A. Southern blot analysis.

Genomic DNA extracted from the purified *E. chaffeensis* (200 ng each) was digested with restriction endonucleases, electrophoresed, and transferred to Hybond-N^+ nylon membrane (Amersham, Arlington Heights, IL), by a standard method. The 0.8-kb *p28* gene fragment from the clone pCRII*p28* was labeled with [α-^{32}P]dATP by the random primer method using a kit (Boehringer Mannheim, Indianapolis, IN) and the labeled fragment was used as a DNA probe. Hybridization was performed at 60°C in rapid hybridization buffer (Amersham) for 20 h. The nylon sheet was washed in 0.1 x SSC (1 x SSC containing 0.15M sodium chloride and 0.015M sodium citrate) with 1% SDS at 55°C and the hybridized probes were exposed to Hyperfilm (Amersham) at -80°C.

Genomic Southern blot analysis with several restriction enzymes resulted in one or more DNA fragment(s) of *E. chaffeensis* which hybridized to ^{32}P-labeled *omp-1* gene probe. The restriction enzymes used did not cut within the *p28* gene portion of the pCRII*p28* insert. *Xba* I, *Bgl* II, and *Kpn* I produced two bands, *Spe* I generated three bands, and *Eco*R V and *Pst* I produced multiple bands with different densities. *Eco*R I generated a broad band of 2.5 to 4kb. These homologous genes are designated as *omp-1* (outer membrane protein-1) family.

B. Cloning and sequencing of genomic copies of *E. chaffeensis omp-1* gene.

The *EcoR* I and *Pst* I fragments of DNA, detected by genomic Southern blot analysis as described above, were inserted into pBluescript II KS (+) vectors, and the recombinant plasmids were introduced into *E. coli* DH5α. Using the colony hybridization method with the ^{32}P-labeled *omp-1* gene probe, four positive clones were isolated from the transformant. The positive clones were designated pEC2.6, pEC3.6, pPS2.6, and pPS3.6. These contained the ehrlichial DNA fragments of 2.6-kb (*Eco*R I), 3.6 kb (*Eco*R I), 2.6 kb (*Pst* I), and 3.6 kb (*Pst* I), respectively. The inserts of the clones pEC3.6 and pPS2.6 overlapped as shown in Fig. 2. The overlapping area was further confirmed by PCR of *E. chaffeensis* genomic DNA with two pairs of primer sets interposing the junctions of the four clones. The 1.1- to 1.6-kb DNA fragments of *Hin*dIII-*Hin*dIII, *Hin*dIII-*Eco*RI, or *Xho*I-*Eco*RI in the pEC2.6 and pEC3.6 were subcloned for sequencing. DNA sequencing was performed with suitable synthetic primers by dideoxy-termination method as described above.

Four DNA fragments from 2.6 to 3.6 kb were cloned from the *Eco*RI-digested and the *Pst*I-digested genomic DNA of *E. chaffeensis* by colony hybridization with radiolabeled *omp-1* gene probe. The inserted DNA of the two recombinant clones, pEC3.6 and PPS2.6, were overlapped. Sequencing revealed one 5'-truncated ORF of 243 bp (designated *omp-1A*) and five complete ORF of 836-861 bp (designated *omp-1B* to *omp-1F*), which are tandemly-arrayed and are homologous to the *p28* gene (but are not identical), in the ehrlichial genomic DNA of 6,292 bp. The intergenic spaces were 581 bp between *omp-1A* and *omp-1B* and 260-308 bp among others. Putative promoter regions and ribosome-binding sites were identified in the noncoding regions.

C. Sequence analysis and GenBank accession number.

Nucleotide sequences were analyzed with the DNASIS program (Hitachi Software Engineering Co., Ltd., Yokohama, Japan). A homology search was carried out with databases of the GenBank, Swiss Plot, PDB and PIR by using the software basic local alignment search tool in the BLAST network service (the National Center for Biotechnology Information, Bethesda, MD). Phylogenetic analysis was performed by using the PHYLIP software package (version 3.5). An evolutional distance matrix, generated by using the Kimura formula in the PROTDIST, was used for construction of a phylogenetic tree by using the unweighted pair-group method analysis (UPGMA) (Felsenstein, J. 1989. PHYLIP-phylogeny inference package (version 3.3). Cladistics 5:164-166). The data were also examined using parsimony analysis (PROTPARS in PHYLIP). A bootstrap analysis was carried out to investigate the stability of randomly generated trees by using SEQBOOT and CONSENSE in the same package. The nucleotide sequence of the

p28 gene and its gene copies has been assigned GenBank accession numbers U72291 and AF021338, respectively.

Proteins encoded by the *omp-1* genes.

Five complete *omp-1* gene copies (*omp-1B* to *omp-1*F) encode 279 to 287-amino acid proteins with molecular masses of 30,320 - 31,508 Da. The 25-amino acid sequence at the N-terminus of OMP-1B to OMP-1F (encoded in *omp-1B* to *omp-1F*) is predicted to be a signal peptide because three carboxyl-terminal amino acids of the signal peptides (Ser-X-Ala in OMP-1B, Leu-X-Ser for OMP-C, and Ser-X-Ser for OMP-1D and OMP-1F) are included in the preferred amino acid sequence of signal peptidase at the processing sites proposed by Oliver. The calculated molecular masses of the mature OMP-1B to OMP-1F from the predicted amino acid sequences are 28,181 Da for OMP-1B, 27,581 Da for OMP-1C, 28,747 Da for OMP-1D, 27,776 Da for OMP-1E, and 27,933 Da for OMP-1F. The estimated isoelectric points are 4.76-5.76 in the mature OMP-1B to OMP-1F. An amino acid sequence in *omp-1F* gene (the 80th to 94th amino acids) was identical to the N-terminal amino acid sequences of *E. chaffeensis* native P23 protein as determined chemically, which indicates that P23 is derived from the *omp-1F* gene.

Alignment of predicted amino acid sequences of the *E. chaffeensis* OMP-1 family and *Cowdria ruminantium*, revealed substitutions or deletions of one or several contiguous amino acid residues throughout the molecules. The significant differences in sequences among the aligned proteins are seen in the regions indicated SV (semivariable region) and HV (hypervariable region) 1 to 3 in Fig 34. Computer analysis for hydropathy revealed that protein molecules predicted from all *omp-1* gene copies contain alternative hydrophilic and hydrophobic motifs which are characteristic of transmembrane proteins. The HV1 and HV2 were found to locate in the hydrophilic regions.

The amino acid sequences of 5 mature proteins without signal peptides (OMP-1, and OMP-1C to OMP-1F) were similar to one another (71-83%) but the sequence of OMP-1B was dissimilar to those of the 5 proteins (45-48%). The amino acid sequences of the 5 proteins showed an intermediate degree of similarity with that of *C. ruminantium* MAP-1 (59-63%), but the similarity between that of the OMP-1B and the *C. ruminantium* MAP-1 was low (45%). These relations are shown in a phylogenetic tree which was obtained based on the amino acid sequence alignment by UPGMA method in the PHYLIP software package. Three proteins (OMP-1, OMP-1D, and OMP-1F) and two proteins (OMP-1C and OMP-1E) formed two separate clusters. The OMP-1B was located distantly from these two clusters. The *C. ruminantium* MAP-1 was positioned between the OMP-1B and other members in the OMP-1 family.

Preparation of a Recombinant form of OMP-1 and P30

The 0.8-kb *p28* gene from *E. chaffeensis* was excised from the clone pCRII*p28* by *Eco*RI-*Not*I double-digestion, ligated into *Eco*RI-*Not*I sites of a pET 29a expression vector, and amplified in *Escherichia coli* BL21 (DE3)pLysS (Novagen, Inc., Madison, WI). The clone (designated pET29*p28*) produced a fusion protein with a 35-amino acid sequence carried from the vector at the N terminus. The amino acid sequence of the OMP-1 portion of the fusion protein, referred to hereinafter as rOMP-1, is depicted in Fig. 1.

An expression vector comprising the *p30* gene was used to prepare the recombinant form of P30. To prepare the expression vector, an 0.6-kb fragment was excised from the clone pCRII*p30* by *Eco*RI digestion, ligated into *Eco*RI site of a pET29a expression vector, and amplifed in *E. coli* BL21(DE3)pLys (Novagen, Inc., Madison, Wis.). The clone (designated pET29*p30*) produced a fusion protein with a 35-amino-acid sequence and a 21-amino-acid sequence carried from the vector at the N and C termini, respectively. The fusion protein had an amino acid sequence consisting of 249-amino acid residues with a molecular mass of 27,316 Da. The amino acid sequence of the P30 portion of the fusion protein, referred to hereinafter as rP30, is amino acid 33 through amino acid 224 of the sequence shown in Fig. 19B.

Preparation of anti-rOMP1 antibody

An rOMP-1 antigen was prepared by excising the gel band corresponding to the rOMP-1 protein in SDS-PAGE, mincing the band in phosphate-buffered saline (PBS), pH 7.4, and mixing with an equal volume of Freund's incomplete adjuvant (Sigma). The rOMP-1 mixture (1 mg of protein each time) was subcutaneously injected into a rabbit every 2 weeks four times. A serum sample was collected from the rabbit to provide the anti-rOMP-1 antibody

The anti-rOMP-1 antibody was examined by western immunoblot analysis. The results indicated that the rabbit anti-rOMP-1 antibody recognized not only rOMP-1 (31 kDa) and OMP-1 protein, but also P29 and P25 of *E. chaffeensis* and P30 of *E. canis* . These results indicate that OMP-1 shares antigenic epitopes with P25 and P29 in *E. chaffeensis* and P30 of *E. canis*.

The following examples are for purposes of illustration only and are not intended to limit the scope of the claims which are appended hereto.

Example 1. Assaying for the presence of anti-OMP-1 antibody in a Patient

Convalescent-phase serum from a patient with clinical signs of human ehrlichiosis was used. Western blot analyses using the rP28 protein as antigen was performed with 1:1,000 dilutions of this serum. Alkaline phosphatase-conjugated affinity-purified anti-human immunoglobulin G (Kirkegaard & Perry Laboratories, Inc., Gaithersburg, MD) was used at a

1:1,000 or 1:2,000 dilution as secondary antibodies. Results indicated that serum from a patient with clinical signs of human ehrlichiosis reacted strongly to rOMP-1 protein (31 kDa).

Example 2 Assaying for the presence of anti-OMP-1 antibody in a Patient

Convalescent-phase serum from a patient with clinical signs of human ehrlichiosis was reacted with the rP30 protein of *E. canis* as described in Example 1. The serum reacted strongly to rP30. These results indicate the rP30 is useful for diagnosing an infection with *E. chaffeensis* in human patients.

Example 3. Identifying *E. chaffeensis*-infected cells using anti-rOMP-1 antibody

E. chaffeensis-infected DH82 cells were sonicated and centrifuged at 400 x g for 10 min. The supernatant was then centrifuged at 10,000 x g for 10 min to obtain ehrlichia-enriched pellet. The pellet was resuspended and incubated with rabbit anti-rOMP-1 antibody or normal rabbit serum (1:100 dilution) at 37°C for 1h in PBS containing 1% bovine serum albumin (BSA-PBS). After washing, the ehrlichiae was incubated with gold-conjugated protein G (20 nm), Sigma) at 1:30 dilution for 1 h at room temperature in BSA-PBS. After washing again, the specimen was fixed with 1.25% formaldehyde, 2.5% glutaraldehyde, and 0.03% trinitrophenol in 0.1 M cacodylate buffer (pH 7.4) for 24h and postfixed in 1% osmium-1.5% potassium ferricyanide for 1 h (34). The section was then embedded in PolyBed 812 (Polysciences, Warraington, Pa). The specimen was ultrathin sectioned at 60 nm, stained with uranyl acetate and lead citrate, and observed with a Philips 300 transmission electron microscope at 60 kV.

Transmission immunoelectron microscopy with colloidal gold-conjugated protein G and rabbit anti-rP28 antibody revealed gold particles bound to *E. chaffeensis* surface. The distribution of the particles was random, close to the surface, and appeared as if almost embedded in the membrane, suggesting that the antigenic epitope protrudes very little from the lipid bilayer. Nonetheless, the antigenic epitope was surface-exposed, and thus, could be recognized by rabbit anti-rOMP-1 antibody. No gold particles were observed on host cytoplasmic membrane or *E. chaffeensis* incubated with normal rabbit serum.

Example 4. Immunization of mice and *E. chaffeensis challenge*.

The rOMP-1 band in SDS- PAGE was excised, minced, and mixed with an equal volume of Freund's incomplete or complete adjuvant. Nine BALB/c male mice (6 weeks old) were divided into two groups. Five mice were intraperitoneally immunized a total of four times at 10-day intervals; twice with a mixture of the minced gel with the rOMP-1 (30 to 40 μg of protein per mouse each time) and incomplete adjuvant, and twice with a mixture of the recombinant protein (the same amount as before) and complete adjuvant. Four mice were intraperitoneally injected with a mixture of the minced gel without protein and the respective adjuvants. For ehrlichia-

challenge, approximately 1 x 10^7 DH82 cells heavily-infected with *E. chaffeensis* were disrupted by sonication in serum-free DMEM (GIBCO-BRL) and centrifuged at 200 x g for 5 min. The supernatant was diluted to a final volume of 5 ml, and 0.3 ml was inoculated intraperitoneally into each mouse 10 days after the last immunization. Before challenge, all 5-immunized mice had a titer of 1:160 against *E. chaffeensis* antigen by IFA and all 4-nonimmunized mice were negative.

At day 5 post-challenge, approximately 1 ml of blood was collected in an EDTA tube from each mouse and protection was assessed by PCR detection of *E. chaffeensis* 16S rDNA in the buffy coat of the collected blood. *E. chaffeensis* could not be reisolated in cell culture at day 10 postinfection. Day 5 post challenge is the optimum time at which establishment of ehrlichial infection can be examined by PCR without the influence of residual DNA from the ehrlichiae used as the challenge before the spontaneous clearance of organisms take place. The *E. chaffeensis*-specific DNA fragment was observed in all nonimmunized mice but not in any immunized mice, indicating that immunization of rOMP-1 apparently protects mice from ehrlichial infection and indicating that the OMP-1 is a potential protective antigen.

Example 5 Assaying for the presence of anti-P30 antibody in Dogs

The rP30 protein was used as an antigen in a Western immunoblot analysis and dot blot analysis to detect the presence of antibody to *E. canis* in serum from *E. canis* infected dogs. The results of the Western immunoblot analysis indicated that reactivity of the sera with rP30 was stronger than the reactivity that was observed when purified *E. canis* was used as antigen. The results of the dot blot assay indicated that rP30 is a useful and sensitive tool for serodiagnosis of canine ehrlichiosis.